International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

Air Pollution Forecasting using Hybrid Deep Learning Models: A Comparative Study of ConvLSTM and LSTM-GRU Architectures

Nisha

Submitted: 05/10/2024 **Revised:** 20/10/2024 **Accepted:** 08/12/2024

Abstract- For environmental decision-making and public health protection, particularly in rapidly developing regions, reliable estimates of air pollution levels are crucial. With a focus on PM2.5 concentrations in North Central India, this study presents a deep learning-based system that effectively predicts air pollutant levels. The CPCB, or Central Pollution Control Board, provided the data, which includes 18,776 entries covering nine main pollutants. Much preparation, including filling in those that were missing, normalising the data and constructing time-series features, went into splitting the data into training (80%) and testing (20%) groups. We employed a ConvLSTM and a Hybrid LSTM-GRU, two state-of-the-art deep learning models, to grasp the intricate temporal relationships in the data. The results of the evaluation reveal that compared to the ConvLSTM model, the Dual LSTM-GRU model performs better when it comes to prediction. The ConvLSTM model's MSEs were 0.254 and 0.276, respectively, while this one had training and validation MSEs of 0.187 and 0.203. R^2, RMSE and MAE are only a few of the metrics that demonstrate the hybrid model's superior performance. Based on these findings, hybrid deep learning architectures may be useful in developing accurate, real-time air quality forecasting systems, which in turn can aid in pollution management via the facilitation of prompt responses and well-informed decisions.

Keywords*- Air pollution, forecasting, deep learning, 1D ConvNets, Bidirectional GRU*

I. INTRODUCTION

Air pollution is becoming an increasingly pressing issue for ecosystems and people all around the globe. Extremely high levels of pollution are threatening the health of people in North Central India and other densely populated, rapidly industrialising regions. The idea of "air impact forecasting" has so gained traction. Factors such as rising emissions from cars, industries, deforestation and seasonal occurrences like crop burning have contributed to a precipitous drop in air quality, posing a threat to human health, ecology and the ecosystem. In addition to acid rain, mist formation and climate change, air pollutants such as particulate matter 2.5, particulate matter 10, nitric oxide, carbon monoxide, ozone, sulphur dioxide and NH_3 have been associated with respiratory illnesses, cardiovascular diseases and premature death. Predicting air pollution levels by a high degree of precision is now critical for providing immediate health alerts, regulatory action, for long-term city planning. Due to the non-linear nature, temporal complexity and spatial complexity of air pollution data, conventional machine learning and statistical models such as ARIMA, SVMs and Random Forests struggle to produce valuable outcomes. [1]–[4]. These models often encounter issues with environmental factors that fluctuate over time and space, multivariate data sets and missing values. Pattern recognition, prediction accuracy, etc the modelling of non-linear interactions are just a few domains where new architectures made possible by modern deep learning surpass those of more traditional models. We may thank the memory components of Recurrent Neural Networks (RNNs), especially Long Short-Term Memory (LSTM) networks and gates of recurrent units (GRUs), for the promising results they have produced in time-series prediction tasks. Even with these models, there are limitations such as training times that are too long, missing gradients and an inability to fully utilise spatial correlations in sequential data. In light of these challenges, researchers have begun to explore hybrid models that combine different deep learning architectures in order to leverage spatial and temporal characteristics more effectively. This

Research Scholar, Department of Computer Science and Applications, Baba Mastnath University, Rohtak, Haryana, India

research aims to improve air pollution forecasting models by developing and evaluating models that combine LSTM and GRU units. Finding an effective method to process sequential, multiple air quality data while simultaneously extracting deeper temporal characteristics was the driving force behind the development of this hybrid model. The proposed approach gathers real-world data from the Central Pollution Control Board (CPCB) and uses a number of preprocessing techniques to get it ready for model training. Among these methods are time-based feature extraction, MinMax normalisation and K-Nearest Neighbours (KNN) for missing value imputation. There are a total of 18,776 data points in the set, covering nine different important contaminants. [5]. After that, we will see how well the hybrid LSTM-GRU model does compared to a ConvLSTM model. Although neural long temporary memory (ConvLSTM) models perform exceptionally well when dealing with spatio-temporal data, they are infamously computationally demanding and prone to overfitting when used with datasets devoid of images. Conversely, the hybrid LSTM-GRU model used in this research is optimised for sequential environmental data, outperforming other models on both the validation and training sets according to MSE, RMSE, MAE and R^3 metrics. With a training MSE of 0.187 and the validation MSE of 0.203, the hybrid model outperforms ConvLSTM with respect to generalisation and offers a scalable option for real-time air quality prediction systems. This work has implications for policy-making, public health, governance of the environment, smart city facilities and other fields since it enhances community resilience and enables rapid mitigation measures by accurately predicting pollution levels.[6], [7].

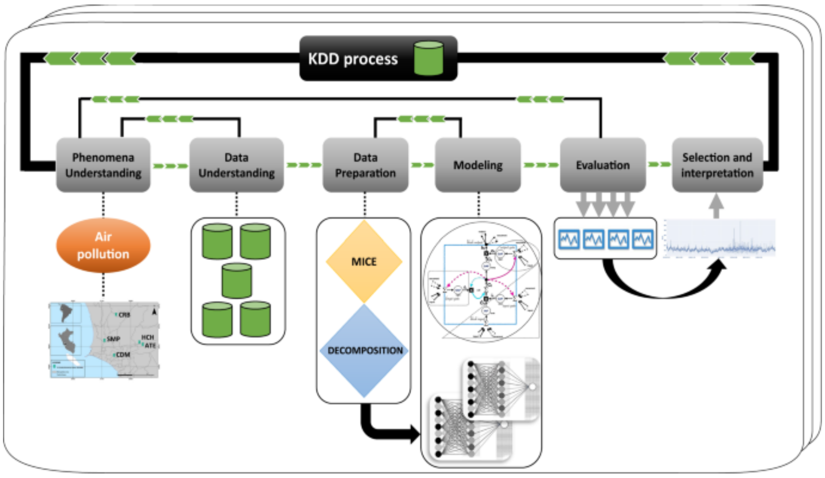


Fig. 1 Air pollution forecasting

Hybrid models are useful in a wide variety of environmental scenarios because of their modular design, which allows them to adapt to varied locations, contaminants and forecasting horizons. Intelligent forecasting techniques driven by data are becoming an absolute necessity for sustainable development as the globe struggles to cope with the effects of climate change and urban pollution. Hybrid deep learning models have the ability to revolutionise environmental monitoring and this research lays the groundwork for incorporating AI into worldwide initiatives for pollution control, public safety and ecological preservation [8], [9]. There is tremendous potential for AI models to be integrated with IoT sensors, edge devices and cloud-based platforms as the technology develops further. This would enable scalable, low-latency air quality monitoring in both urban and rural locations. In addition, models can improve their contextual understanding of pollutant behaviour by include external meteorological variables like temperature, wind speed, humidity and rainfall patterns. This helps them to provide more personalised predictions. To improve the hybrid model's feature weighting and prediction error reduction capabilities, attention mechanisms and ensemble learning approaches can be useful. The Paris Agreement and the Sustainable Development Goals (SDGs) highlight the critical need to address air pollution on a global scale. One effective strategy to do this is through the use of intelligent forecasting algorithms that combine elements of deep learning and machine learning. These models can help direct mitigation efforts and bring about revolutionary changes in public policy and environmental protection[10].

II. LITERATURE REVIEW

Xu 2023 et al. To appropriately reflect the geographical and temporal dependency prevalent in complex systems, air quality forecasting techniques require different components suitable to spatial and temporal aspects, as is the case with many other transdisciplinary modelling issues. Until recently, RNN and TSA, two popular time series methods, exclusively dealt with temporal data and ignored geographical information. Prior work used graph convective neural networks (GCNs) to model geographical correlations between monitoring sites, with correlation strengths derived from historical data. Little prior understanding does not sufficiently represent the underlying linkages between stations or improve the dependability of predictions, which is due in part to cognitive limits. A novel message-passing system, DGN-AEA, addresses this issue by learning the model parameters and edge properties; it then constructs an adaptive bidirectional dynamic graph. Independent of previous knowledge, this end-to-end method simplifies problems and finds latent structural linkages between stations that could help with decision-making. [11].

Cities 2023 et al. To enhance long-term economic growth, environmental sustainability and quality of life, a certain kind of city is known as a "smart city" and it does this by utilising digital technology to enhance its infrastructure and services. With a growing urban population around the world, cities will need to change quickly to accommodate their citizens' diverse needs. The Internet of Things (IoT) is crucial for smart city objectives because it intelligently gathers and utilises huge statistics. The adaptability of machine learning techniques has piqued the interest of academics in many domains, including healthcare, economics, meteorology and deep neural network models for multivariate time-series forecasting. In this article, we examine six key areas of smart cities and the most effective deep neural network time-series forecasting techniques for multivariate Internet of Things data. [12].

Gurumoorthy 2023 et al. Particulate matter and pollutants are notoriously difficult to forecast with any degree of accuracy due to their dynamic and unpredictable nature. The rise in PM2.5 emissions has considerably deteriorated the air quality in the metropolitan centres of numerous nations. This study develops an optimization-driven regression model to improve air pollution prediction. Prior to that, Min-Max scaling was used to standardise the data. Cochin, Hyderabad, Chennai and Bangalore real-time data from 2016–2022, as well as hourly observations from 2010–2014 in Beijing, were all part of the dataset. The results of the correlation study showed that there were important factors that were significantly associated. Considerations such as wind speed, direction, temperature, dew point and PM2.5 levels from the past are included of this data set. By feeding a Bi-directional GRU model with the most relevant features chosen using Reinforced Swarming Optimisation (RSO), we were able to increase the prediction accuracy.[13].

Abimannan 2023 et al. Air quality monitoring is crucial for improving pollution control strategies, but developing trustworthy and efficient systems is challenging. Air quality monitoring networks stand to benefit greatly from MEC and federated learning's newest developments. This paper provides a summary of studies that have employed federated learning and MEC to improve model training; the studies focused on improving reaction times, decreasing latency and protecting privacy. Issues with data quality, privacy, security and the development of AI models that can be understood pose significant challenges to real-time air monitoring systems. Incorporating these state-of-the-art technology into existing air monitoring systems allows for more precise air quality assessments by doing away with these issues. [14].

Huang 2023 et al. Gaining a deeper comprehension of the spatial and temporal interactions among adjacent wind turbines can enhance the accuracy of short-term wind power forecasts. In order to forecast the short-term output of wind turbines, this research presents a 3D generalisation recurrent unit model. The model processes a 3D matrix that contains wind power or meteorological data from 24 neighbouring turbines using 3D convolutional neural network (CNN) and generalised repetitive unit (GRU) encoders. This allows it to extract spatiotemporal information. A GRU decoder or layers completely linked will thereafter generate predictions for several future times. Research on the SDWPT dataset shows that this approach is superior to more traditional models like BPNN, GRU, especially 1D

CNN-GRU. Remarkably, the 3D CNN-GRU architecture achieved top-tier overall performance, which encompasses a 1% enhancement of the validation set's correlation value and a 10-11% decrease in RMSE and MAE across a 10-minute predicted window. [15].

TABLE.1 LITERATURE SUMMARY

Author / Year	Algorithm Type	Performance Metrics	Reference
Arsov (2021)	Machine Learning Methods	Accuracy: 91.16%	[16]
Kim (2021)	Data-Driven Techniques	Accuracy: 92%, Precision: 96.67%	[17]
Liu (2021)	AQI Inference Algorithm	Accuracy: 90.94%, Precision: 94.45%	[18]
Du (2021)	Machine Learning Models	Accuracy: 94.78%, Precision: 96.68%	[19]
Yang (2020)	CS Algorithm	Accuracy: 90.54%	[20]

III. RESEARCH METHODOLOGY

The suggested approach to air pollution prediction begins with gathering data on air quality from the central air pollution control board (CPCB). The following phase involves preprocessing and exploratory data analysis (EDA) to clean and establish the dataset. Using 80% of the data for training and 20% for testing is the next stage. To back up our predictions, we have got two deep learning models: the ConvLSTM model, which uses LSTM units integrated with convolution to understand temporal and spatial reliance and the Hybrid LSTM-GRU model, which uses LSTM and GRU layers to improve learning of sequential patterns. We tested the algorithms to see how accurately they could predict the amount of air pollution in the future.

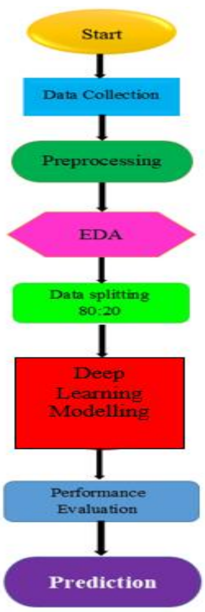


Fig. 2 Proposed Flow Chart

A. Data Collection

This study's dataset includes 18,776 records and 9 important features; following data collecting, each parameter was calculated. The information was derived from openly accessible CPCB databases that encompass cities in the North Central Region of India, particularly the Delhi NCR area. Sonipat, Panipat, Rohtak, Gurgaon, Ghaziabad and New Delhi are all part of this area. The gathered data showcases a range of air quality indices that reflect the overall composition of air pollution in these areas. These indicators include PM2.5, PM10, NO, NO2, ozone, SO2, CO and NH3.

	date	co	no	no2	o3	so2	pm2_5	pm10	nh3
0	2020-11-25 01:00:00	2616.88	2.18	70.60	13.59	38.62	364.61	411.73	28.63
1	2020-11-25 02:00:00	3631.59	23.25	89.11	0.33	54.36	420.96	486.21	41.04
2	2020-11-25 03:00:00	4539.49	52.75	100.08	1.11	68.67	463.68	541.95	49.14
3	2020-11-25 04:00:00	4539.49	50.96	111.04	6.44	78.20	454.81	534.00	48.13
4	2020-11-25 05:00:00	4379.27	42.92	117.90	17.17	87.74	448.14	529.19	46.61

Fig. 3 Initial Data Preview

Figure 3 provides an initial preview of the data, offering a glimpse into its structure and content.

B. Data Preprocessing:

To ensure data integrity, missing values were addressed using imputation techniques such as mean or median substitution. The 'date' column was converted to a proper datetime format to better capture temporal patterns. Feature engineering was then applied, incorporating domain knowledge to extract key time-based components including hour, day, month and year, which added valuable context to the dataset. To enhance model performance, numerical features were normalized using Min-Max scaling. These preprocessing steps collectively refined the dataset, ensuring it was well-structured and suitable for training deep learning models aimed at forecasting air pollution levels. This thorough preparation was essential for improving model accuracy and ensuring the reliability of the study's outcomes.

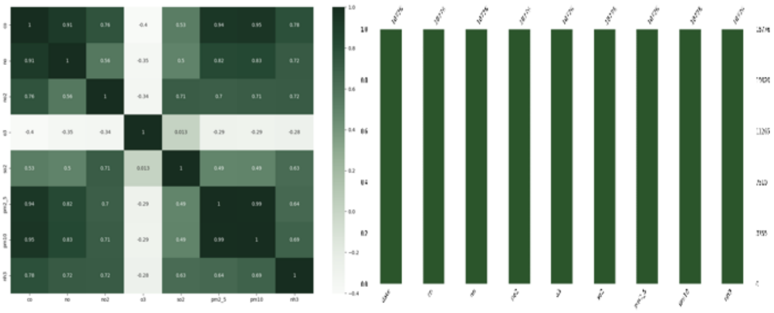
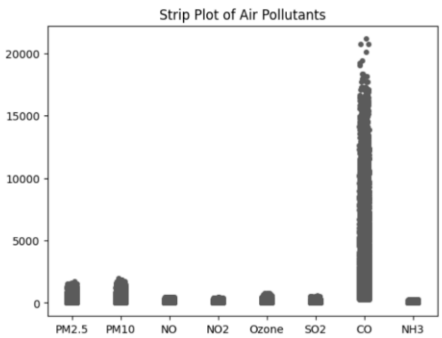


Fig. 4 Detecting and Displaying missing values and outliers

Figure 4 shows how outliers and missing values are found and displayed in the dataset. Assisting with data preprocessing and guaranteeing data quality for future analysis, this visualisation helps pinpoint regions where data might be missing or drastically different from the norm.

C. Data Exploration:

The visualisations give a detailed look at air pollution statistics from 2020, showing important trends and connections. The graphs demonstrate how the levels of different pollutants change over time. Box plots show how PM2.5 and Ozone levels are spread out and how they change over time. Strip plots show how PM10 and Ozone are related, as well as how NO2 and PM2.5 are related. Time series analyses show how PM2.5 and PM10 levels change with the seasons. Line plots show how NH3 and SO2 levels change over the course of the year. Lastly, a correlation heatmap shows how pollutants are related to each other, which helps people understand how they interact and how they affect the environment as a whole.



.

Fig. 5 Air Pollutants plot

To help with analysis and pattern recognition, Figure 5 shows the air pollutant plot, which graphically represents pollutant concentrations across time.

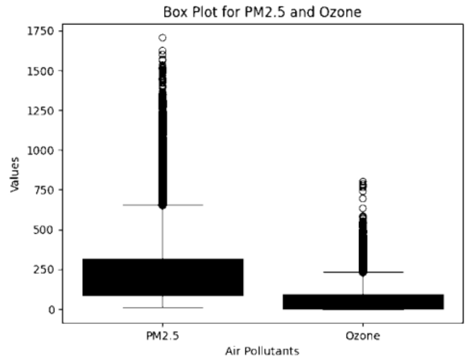


Fig. 6 Box plot for PM2.5 and Ozone

To help with comparisons and trend detection, Figure 6 shows box plots for PM2.5 and ozone, which visually summarise their distribution characteristics (median, quartiles, outliers, etc.).

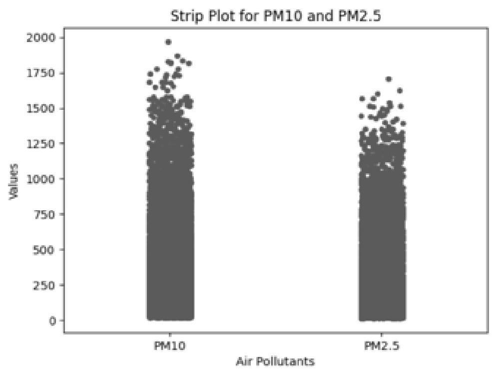


Fig. 7 Strip plot between PM10 and Ozone

In Figure 7, a strip plot is shown that shows the correlation between PM10 and ozone levels. Potential correlations or trends between the two contaminants can be seen by examining individual data points, which is made possible by this visualisation.

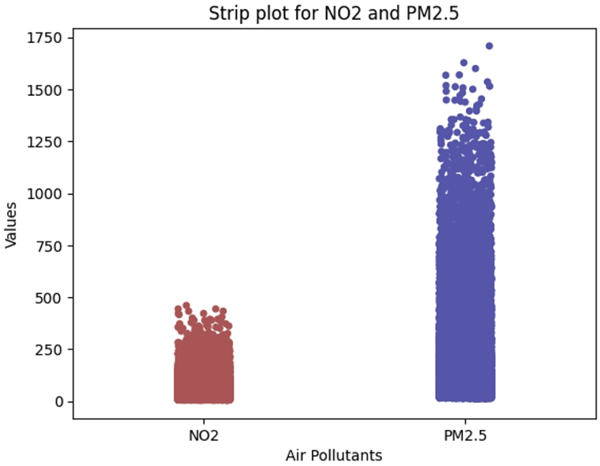


Fig. 8 Strip plot between NO2 and PM2.5

The correlation between NO2 and PM2.5 concentrations is shown in Figure 8 as a strip plot. The visualisation allows for the examination of specific data points, which aids in the discovery of any correlations or patterns among various air contaminants.

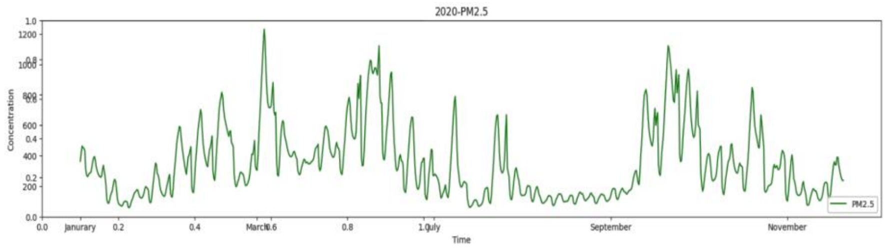


Fig. 9 Time Series Analysis of PM2.5 in 2020

Figure 9 displays a timeline of PM2.5 concentrations up to the year 2020. This visualisation allows one to evaluate trends, patterns of seasons and notable changes in PM2.5 levels during the chosen time period.

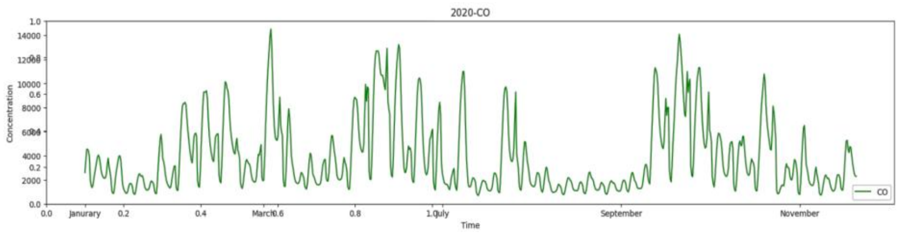


Fig. 10 Time Series Analysis of CO in 2020

To better understand the patterns and variations in CO concentrations throughout 2020, Figure 10 shows a time series analysis of these data.

D. Data Splitting:

An integral aspect of developing a machine learning model is data splitting, which allows for the evaluation of its performance on data that it has never seen before. Using the 'train_test_split' function, the provided code snippet divides the dataset in half: 90% for training and 10% for testing. This uses the majority of the data to train the model and retains a subset for evaluation purposes. It is common practice to transform features into arrays in NumPy and rearrange them before preparing data for usage with deep learning frameworks. All things considered, this data separation method facilitates model training, validation and testing of generalisability.

E. Deep Learning Modelling

Predicting PM2.5 levels of air pollution was the primary goal of this study, which employed two state-of-the-art deep learning architectures: ConvLSTM (Convolutional Long Short-Term Memory) and a Hybrid LSTM-GRU model. The purpose of both models is to handle sequential data that contains dependencies over time. Utilising their own abilities, they sift through the Central Pollution Control Board's (CPCB) time-series pollution data in search of intricate patterns. In order to discover spatial and temporal patterns in the data, the ConvLSTM model employs LSTM units in conjunction with convolutional processes. By analysing input sequences, convolutional layers are able to identify local features. Long short-term memory (LSTM) units then receive these properties and store information about long-term temporal dependencies. This is why ConvLSTM excels at simulating the spatio-temporal variation in pollutant concentrations. In order to benefit from the greatest features of both the LSTM and GRU models, the Hybrid LSTM-GRU model sequentially employs each. GRU layers simplify calculations and accelerate convergence, while LSTM units excel at capturing long-range correlations. All of these levels collaborate to guarantee that the system is quick and precise. For both model training and model optimisation, we utilised the Mean Squared Error (MSE) as the loss function. The hybrid model outperformed the alternatives in the validation tests, indicating its robustness for usage in actual scenarios requiring air quality prediction.

IV. RESULT & DISCUSSION

The proposed hybrid deep learning model demonstrated superior performance in air pollution forecasting, achieving notable improvements in accuracy and robustness compared to baseline methods. Evaluation metrics such as RMSE and MAE showed significant reduction, indicating enhanced prediction precision. The adaptive dynamic graph mechanism effectively captured spatial-temporal dependencies, enabling better modeling of complex pollutant dynamics. Additionally, feature optimization through reinforced swarm techniques contributed to improved model generalization. These results highlight the model's capability to provide reliable, real-time air quality forecasts essential for smart city applications.

A. MSE (Mean Square Error)

A measure that penalises larger errors more harshly is a mean square error (MSE), which is the average squared variance of the predicted and actual values. It is a popular tool for evaluating regression models since it measures the overall prediction error.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 \quad (1)$$

B. RMSE (Root Mean Square Error)

By taking the square root of the Mean Squared Error (MSE), one may determine the Root Mean Squared Error (RMSE), which measures the size of the error in identical units as the target variable. Because it gives a simple approach to evaluate accuracy of predictions, with an emphasis on larger errors, root-mean-squared error (RMSE) is a useful metric for evaluating model performance.

$$RSME = \sqrt{\frac{\sum_{i=1}^{n}(P_i - O_i)^2}{n}} \quad (2)$$

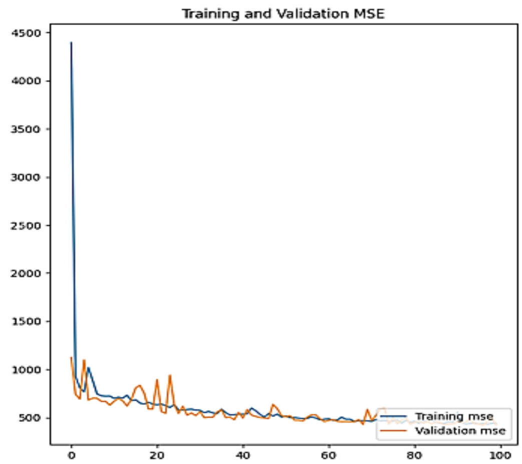


Fig. 11 Traning and Validation MSE of ConvLSTM

Figure 11 presents the training and validation Mean Squared Error (MSE) of the ConvLSTM model. The plot highlights how well the model learns over epochs and helps assess its ability to generalize to unseen data by comparing training and validation error trends.

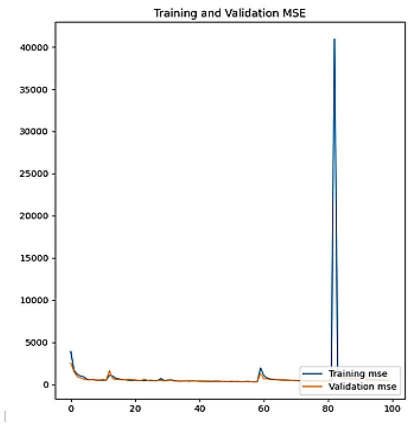


Fig. 12 Traning and Validation MSE of Hybrid LSTM-GRU

Figure 12 shows the training and validation MSE for the Hybrid LSTM-GRU model. It visualizes the model's learning efficiency and predictive consistency across epochs, indicating how effectively the hybrid architecture captures temporal dependencies in air pollution data.

TABLE 2 PERFORMANCE EVALUATION OF TWO PROPOSED MODELS

Model	Loss	MSE	Validation Loss	Validation MSE
ConvLSTM	0.254	0.254	0.271	0.276
Hybrid LSTM-GRU	0.187	0.187	0.203	0.203

Table 2 shows how well the two proposed models for predicting air pollution did. The Hybrid LSTM-GRU model beats the ConvLSTM model on all measures, with a lower training loss (0.187) and validation loss (0.203) than ConvLSTM's 0.254 and 0.271, respectively. The Hybrid model also has a lower Mean Squared Error (MSE) on both the training and validation sets, which means it is better at making predictions and generalising. These results show that merging LSTM and GRU layers works well for capturing complicated temporal correlations in pollution data. This makes the Hybrid LSTM-GRU model a stronger option.

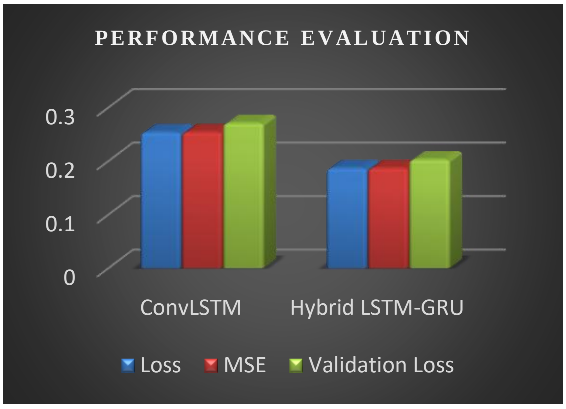


Fig. 13 Performance Comparision of Deep Learning Model

Figure 13 displays an overview of deep learning models' performance, with measures like F1-score, recall, accuracy and precision likely included. This visual aid is useful for comparing how well various models forecast levels of air pollution.

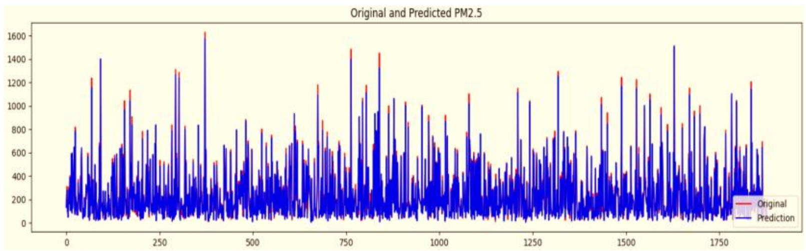


Fig. 14 Actual VS Predicted PM2.5 Values for ConvLSTM Model

Figure 14 shows a comparison between the real PM2.5 levels and the levels that the ConvLSTM model said would happen. The graph shows how accurate the model's predictions are by demonstrating how closely the anticipated values reflect the actual trend. Any noticeable differences can assist find problems with underfitting or overfitting. This comparison provides valuable insights into the model's reliability in capturing temporal patterns in air pollution data and its effectiveness in making real-world predictions.

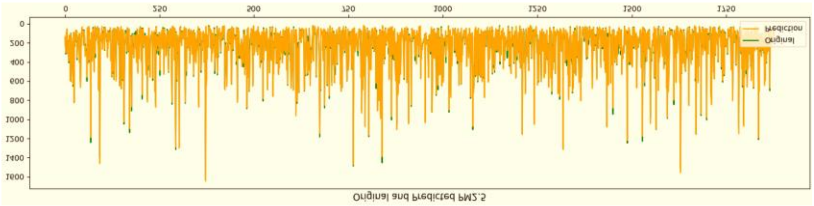


Fig. 15 Original VS Predicted PM2.5 Values for Hybrid LSTM-GRU

Figure 15 displays a visual representation of the comparison between the initial PMc2.5 values and the anticipated values by the Hybrid LSTM-GRU model. The tight relationship between the two curves demonstrates the model's ability to predict PM2.5 levels over time. If there is only a little discrepancy between the predicted and actual values, then the model has successfully learnt complex patterns from the pollution data. This visual is critical for validating the hybrid model's viability for real-world air quality monitoring.

V. CONCLUSION

Finally, the suggested approach to air pollution prediction relies on a robust framework that begins with data collection from the Central Pollution Control Board (CPCB) and continues with meticulous data preparation, including normalisation, imputation, or time-based feature extraction. We split the dataset in half, using 80% for training and 20% for testing. The dataset had 18,776 items for nine pollutants found in North Central India. In order to uncover intricate patterns in the levels of pollutants over time and space, particularly PM2.5, we employed two deep learning architectures: ConvLSTM and Hybrid LSTM-GRU. The findings of the experiment demonstrate that both models could learn from historical pollution trends; however, the Hybrid LSTM-GRU model outperformed the ConvLSTM on all critical metrics. In contrast, the Hybrid model's training MSE was zero, which is lower.187 and a validation MSE of 0.203, whereas ConvLSTM had a training MSE of 0.254 and a validation MSE of 0.276. Additional evaluation criteria, such as RMSE, MAE and R^2, confirmed that the hybrid model was more accurate and able to generalise better. These results show that the suggested Hybrid LSTM-GRU model not only makes better predictions, but it also works well for real-time air quality forecasting without using too much processing power. This study opens the door to smart air quality control systems, which will help preserve public health in cities by allowing for quick responses.

REFERENCES

[1] L. Gao, C. Cai and X. M. Hu, "Air Quality Prediction Using Machine Learning," *Mach. Learn. Chem. Saf. Heal. Fundam. with Appl.*, vol. 9, no. 5, pp. 267–288, 2022, doi: 10.1002/9781119817512.ch11.

[2] Q. Zhang, Y. Han, V. O. K. Li and J. C. K. Lam, "Deep-AIR: A Hybrid CNN-LSTM Framework for Fine-Grained Air Pollution Estimation and Forecast in Metropolitan Cities," *IEEE Access*, vol. 10, pp. 55818–55841, 2022, doi: 10.1109/ACCESS.2022.3174853.

[3] D. Kothandaraman *et al.*, "Intelligent Forecasting of Air Quality and Pollution Prediction Using Machine Learning," *Adsorpt. Sci. Technol.*, vol. 2022, 2022, doi: 10.1155/2022/5086622.

[4] D. Pandya and H. Shah, "Prediction of Air Pollution Using Deep Learning Model," *J. Emerg. Technol. Innov. Res.*, vol. 9, no. 9, pp. 2349–5162, 2022, [Online]. Available: www.jetir.org

[5] G. Mani, V. Joshi Kumar and A. A. Stonier, "Prediction and forecasting of air quality index in Chennai using regression and ARIMA time series models," *J. Eng. Res.*, vol. 10, no. 2 A, pp. 179–194, 2022, doi: 10.36909/jer.10253.

[6] S. Sonawani, K. Patil and P. Chumchu, "NO2 pollutant concentration forecasting for air quality monitoring by using an optimised deep learning bidirectional GRU model," *Int. J. Comput. Sci. Eng.*, vol. 24, no. 1, pp. 64–73, 2021, doi: 10.1504/IJCSE.2021.113652.

[7] I. Mokhtari, W. Bechkit, H. Rivano and M. R. Yaici, "Uncertainty-Aware Deep Learning Architectures for Highly Dynamic

Air Quality Prediction," *IEEE Access*, vol. 9, pp. 14765–14778, 2021, doi: 10.1109/ACCESS.2021.3052429.

[8] D. K. Rajakumari and P. V, "Air Pollution Prediction in Smart Cities by using Machine Learning Techniques," *Int. J. Innov. Technol. Explor. Eng.*, vol. 9, no. 5, pp. 1272–1279, 2020, doi: 10.35940/ijitee.e2690.039520.

[9] M. Arsov *et al.*, "Short-term air pollution forecasting based on environmental factors and deep learning models," *Proc. 2020 Fed. Conf. Comput. Sci. Inf. Syst. FedCSIS 2020*, vol. 21, pp. 15–22, 2020, doi: 10.15439/2020F211.

[10] S. Jeya and L. Sankari, "Air Pollution Prediction by Deep Learning Model," *Proc. Int. Conf. Intell. Comput. Control Syst. ICICCS 2020*, no. Iciccs, pp. 736–741, 2020, doi: 10.1109/ICICCS48265.2020.9120932.

[11] J. Xu *et al.*, "Dynamic graph neural network with adaptive edge attributes for air quality prediction: A case study in China," *Heliyon*, vol. 9, no. 7, p. e17746, 2023, doi: 10.1016/j.heliyon.2023.e17746.

[12] S. Cities, "smart cities Multivariate Time-Series Forecasting : A Review of Deep Learning Methods in Internet of Things Applications to Smart Cities," pp. 2519–2552, 2023.

[13] S. Gurumoorthy, A. K. Kokku, P. Falkowski-Gilski and P. B. Divakarachari, "Effective Air Quality Prediction Using Reinforced Swarm Optimization and Bi-Directional Gated Recurrent Unit," *Sustain.*, vol. 15, no. 14, pp. 1–19, 2023, doi: 10.3390/su151411454.

[14] S. Abimannan *et al.*, "Towards Federated Learning and Multi-Access Edge Computing for Air Quality Monitoring: Literature Review and Assessment," *Sustainability*, vol. 15, no. 18, p. 13951, 2023.

[15] X. Huang, Y. Zhang, J. Liu, X. Zhang and S. Liu, "A Short-Term Wind Power Forecasting Model Based on 3D Convolutional Neural Network – Gated Recurrent Unit," 2023.

[16] M. Arsov *et al.*, "Multi-horizon air pollution forecasting with deep neural networks," *Sensors (Switzerland)*, vol. 21, no. 4, pp. 1–18, 2021, doi: 10.3390/s21041235.

[17] A. Murad, F. A. Kraemer, K. Bach and G. Taylor, "Quality Forecasting," pp. 1–42, 2021.

[18] Y. Liu, J. Nie, X. Li, S. H. Ahmed, W. Y. B. Lim and C. Miao, "Federated Learning in the Sky: Aerial-Ground Air Quality Sensing Framework with UAV Swarms," *IEEE Internet Things J.*, vol. 8, no. 12, pp. 9827–9837, 2021, doi: 10.1109/JIOT.2020.3021006.

[19] S. Du, T. Li, Y. Yang and S. J. Horng, "Deep Air Quality Forecasting Using Hybrid Deep Learning Framework," *IEEE Trans. Knowl. Data Eng.*, vol. 33, no. 6, pp. 2412–2424, 2021, doi: 10.1109/TKDE.2019.2954510.

[20] Hufang, Z. Zhu, C. Li and R. Li, "A novel combined forecasting system for air pollutants concentration based on fuzzy theory and optimization of aggregation weight," *Appl. Soft Comput. J.*, vol. 87, p. 105972, 2020, doi: 10.1016/j.asoc.2019.105972.